

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/559,986  
Source: IFwo  
Date Processed by STIC: 9/20/06

***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 09/20/2006

PATENT APPLICATION: US/10/559,986

TIME: 10:53:24

Input Set : E:\PAT 60100W-1 Sequence Listing.txt  
 Output Set: N:\CRF4\09202006\J559986.raw

3 <110> APPLICANT: Nestec S.A.; Cornell Research Foundation, Inc.  
 5 <120> TITLE OF INVENTION: Modulation Of Coffee Flavour Precursor Levels In Green  
 Coffee Grains  
 7 <130> FILE REFERENCE: PAT 60100W-1  
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/559,986  
 C--> 10 <141> CURRENT FILING DATE: 2005-12-08  
 12 <150> PRIOR APPLICATION NUMBER: EP 03394056.0  
 13 <151> PRIOR FILING DATE: 2003-06-20  
 15 <160> NUMBER OF SEQ ID NOS: 16  
 17 <170> SOFTWARE: PatentIn version 3.1  
 20 <210> SEQ ID NO: 1  
 21 <211> LENGTH: 1543  
 22 <212> TYPE: DNA  
 23 <213> ORGANISM: Coffea canephora  
 25 <220> FEATURE:  
 26 <221> NAME/KEY: mRNA  
 27 <222> LOCATION: (1)..(1543)  
 28 <223> OTHER INFORMATION: mRNA  
 30 <220> FEATURE:  
 31 <221> NAME/KEY: CDS  
 32 <222> LOCATION: (122)..(1315)  
 33 <223> OTHER INFORMATION: CDS  
 35 <400> SEQUENCE: 1  
 36 aaggcagtgg aacaacgcag agtacgcggg ggacactcct ccccgttcca ttccagacca 60  
 37 gggtccaaaa ccaccgtcca agagaggagc agactgcaga gtgatacata caggcacaaa 120  
 38 g atg atg atg aca agc gga ggt ctg atg cta acc tgc act ctg gct att 169  
 39 Met Met Met Thr Ser Gly Gly Leu Met Leu Thr Cys Thr Leu Ala Ile  
 40 1 5 10 15  
 41 acc ctc tta tcc tgc gca ctc atc tct tca acc act ttc caa cat gaa 217  
 42 Thr Leu Leu Ser Cys Ala Leu Ile Ser Ser Thr Thr Phe Gln His Glu  
 43 20 25 30  
 44 att cag tat cga gta caa gac ccg tta atg ata cgc caa gtc acc gac 265  
 45 Ile Gln Tyr Arg Val Gln Asp Pro Leu Met Ile Arg Gln Val Thr Asp  
 46 35 40 45  
 47 aat cac cac cac cgc cac ccc ggt agg tct tct gca aac cat cgt 313  
 48 Asn His His Arg His His Pro Gly Arg Ser Ser Ala Asn His Arg  
 49 50 55 60  
 50 cta ctg ggc acc acc aca gag gtt cac ttc aag tcc ttc gtg gag gag 361  
 51 Leu Leu Gly Thr Thr Glu Val His Phe Lys Ser Phe Val Glu Glu  
 52 65 70 75 80  
 53 tac gag aaa act tac tct acg cac gag gag tac gtg cac cgc ctg ggg 409  
 54 Tyr Glu Lys Thr Tyr Ser Thr His Glu Glu Tyr Val His Arg Leu Gly  
 55 85 90 95  
 56 att ttc gcc aag aac ctc atc aag gcc gcg gag cac cag gcc atg gac 457

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/559,986

DATE: 09/20/2006

TIME: 10:53:24

Input Set : E:\PAT 60100W-1 Sequence Listing.txt

Output Set: N:\CRF4\09202006\J559986.raw

57	Ile	Phe	Ala	Lys	Asn	Leu	Ile	Lys	Ala	Ala	Glu	His	Gln	Ala	Met	Asp	
58							100		105						110		
59	ccc	tcc	gca	atc	cac	ggc	gtc	acc	cag	ttc	tct	gat	ctc	acc	gag	gag	505
60	Pro	Ser	Ala	Ile	His	Gly	Val	Thr	Gln	Phe	Ser	Asp	Leu	Thr	Glu	Glu	
61							115		120				125				
62	gag	ttt	gag	gct	acg	tac	atg	ggc	ctt	aaa	ggt	ggc	gct	gga	gtt	ggt	553
63	Glu	Phe	Glu	Ala	Thr	Tyr	Met	Gly	Leu	Lys	Gly	Gly	Ala	Gly	Val	Gly	
64							130		135				140				
65	ggg	acc	acc	cag	ctg	ggg	aaa	gat	gat	ggg	gat	gag	agt	gca	gca	gag	601
66	Gly	Thr	Thr	Gln	Leu	Gly	Lys	Asp	Asp	Gly	Asp	Glu	Ser	Ala	Ala	Glu	
67	145					150				155			160				
68	gtg	atg	atg	gat	gta	tct	gat	ttg	ccg	gag	agt	ttt	gat	tgg	aga	gaa	649
69	Val	Met	Met	Asp	Val	Ser	Asp	Leu	Pro	Glu	Ser	Phe	Asp	Trp	Arg	Glu	
70							165			170			175				
71	aaa	ggt	gct	gtg	acc	gaa	gtg	aag	acg	cag	gga	aga	tgt	gga	tcg	tgt	697
72	Lys	Gly	Ala	Val	Thr	Glu	Val	Lys	Thr	Gln	Gly	Arg	Cys	Gly	Ser	Cys	
73						180			185			190					
74	tgg	gct	ttt	agt	aca	act	gga	gct	att	gaa	gga	gct	aat	ttc	att	gca	745
75	Trp	Ala	Phe	Ser	Thr	Thr	Gly	Ala	Ile	Glu	Gly	Ala	Asn	Phe	Ile	Ala	
76						195			200			205					
77	act	ggc	aag	ctt	ctc	agc	cta	agt	gaa	cag	cag	ott	gtg	gat	tgt	gat	793
78	Thr	Gly	Lys	Leu	Leu	Ser	Leu	Ser	Glu	Gln	Gln	Leu	Val	Asp	Cys	Asp	
79						210			215			220					
80	cat	atg	tgt	gat	tta	aaa	gaa	aaa	gat	gac	tgt	gat	gat	gga	tgc	tcc	841
81	His	Met	Cys	Asp	Leu	Lys	Glu	Lys	Asp	Asp	Cys	Asp	Asp	Gly	Cys	Ser	
82	225					230				235			240				
83	gga	ggg	cta	atg	aca	act	gct	ttc	aac	tac	ttg	ata	gag	gca	gga	ggt	889
84	Gly	Gly	Leu	Met	Thr	Thr	Ala	Phe	Asn	Tyr	Leu	Ile	Glu	Ala	Gly	Gly	
85						245				250			255				
86	ata	gag	gag	gag	gta	acc	tat	ccc	tac	act	ggg	aaa	cgc	gga	gaa	tgc	937
87	Ile	Glu	Glu	Glu	Val	Thr	Tyr	Pro	Tyr	Thr	Gly	Lys	Arg	Gly	Glu	Cys	
88						260			265			270					
89	aaa	ttc	aat	cct	gag	aaa	gtt	gcg	gtg	aaa	gtg	cgg	aat	ttc	gca	aaa	985
90	Lys	Phe	Asn	Pro	Glu	Lys	Val	Ala	Val	Lys	Val	Arg	Asn	Phe	Ala	Lys	
91						275			280			285					
92	atc	cct	gag	gat	gag	agt	caa	att	gct	gcc	aat	gta	gtg	cat	aat	ggc	1033
93	Ile	Pro	Glu	Asp	Glu	Ser	Gln	Ile	Ala	Ala	Asn	Val	Val	His	Asn	Gly	
94						290			295			300					
95	ccg	ctt	gct	att	gga	ttg	aat	gcg	gta	ttc	atg	caa	act	tac	atc	ggg	1081
96	Pro	Leu	Ala	Ile	Gly	Leu	Asn	Ala	Val	Phe	Met	Gln	Thr	Tyr	Ile	Gly	
97	305					310				315			320				
98	ggt	gtg	tca	tgt	cct	ttt	att	tgt	gac	aaa	aag	agg	atc	aac	cat	ggt	1129
99	Gly	Val	Ser	Cys	Pro	Leu	Ile	Cys	Asp	Lys	Lys	Arg	Ile	Asn	His	Gly	
100						325			330			335					
101	gtt	ctt	ctt	gtg	ggc	tat	ggt	tct	aga	ggc	ttc	tca	atc	ctt	agg	ctt	1177
102	Val	Leu	Leu	Val	Gly	Tyr	Gly	Ser	Arg	Gly	Phe	Ser	Ile	Leu	Arg	Leu	
103						340			345			350					
104	ggc	tac	aag	cca	tac	tgg	att	atc	aag	aac	tca	tgg	ggg	aag	cgt	tgg	1225
105	Gly	Tyr	Lys	Pro	Tyr	Trp	Ile	Ile	Lys	Asn	Ser	Trp	Gly	Lys	Arg	Trp	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/559,986

DATE: 09/20/2006

TIME: 10:53:24

Input Set : E:\PAT 60100W-1 Sequence Listing.txt  
 Output Set: N:\CRF4\09202006\J559986.raw

106	355	360	365	
107	ggc gaa cat ggt tgc tac cgg ctt tgt cga ggg cac aac atg tgt gga			1273
108	Gly Glu His Gly Cys Tyr Arg Leu Cys Arg Gly His Asn Met Cys Gly			
109	370	375	380	
110	atg agc aca atg gtt tca gct gtg gtg aca cag acc tct tga			1315
111	Met Ser Thr Met Val Ser Ala Val Val Thr Gln Thr Ser			
112	385	390	395	
113	tacccaaaaca tctctgtct tcagagggtg tataacaaggt ggtttgcctct tgaaagatct			1375
114	tatcatgttt tcgaaatatt taggttgta taatatgaag ggttagagagt aataagaacc			1435
115	aaacaaaagt tcaggcctgt ttctgtatagg aatgaaat gatcgagtc atttgtact			1495
116	ggatcacaaa aaaaaatcca aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa			1543
119	<210> SEQ ID NO: 2			
120	<211> LENGTH: 397			
121	<212> TYPE: PRT			
122	<213> ORGANISM: Coffea canephora			
124	<400> SEQUENCE: 2			
125	Met Met Met Thr Ser Gly Gly Leu Met Leu Thr Cys Thr Leu Ala Ile			
126	1	5	10	15
127	Thr Leu Leu Ser Cys Ala Leu Ile Ser Ser Thr Thr Phe Gln His Glu			
128	20	25	30	
129	Ile Gln Tyr Arg Val Gln Asp Pro Leu Met Ile Arg Gln Val Thr Asp			
130	35	40	45	
131	Asn His His His Arg His His Pro Gly Arg Ser Ser Ala Asn His Arg			
132	50	55	60	
133	Leu Leu Gly Thr Thr Glu Val His Phe Lys Ser Phe Val Glu Glu			
134	65	70	75	80
135	Tyr Glu Lys Thr Tyr Ser Thr His Glu Glu Tyr Val His Arg Leu Gly			
136	85	90	95	
137	Ile Phe Ala Lys Asn Leu Ile Lys Ala Ala Glu His Gln Ala Met Asp			
138	100	105	110	
139	Pro Ser Ala Ile His Gly Val Thr Gln Phe Ser Asp Leu Thr Glu Glu			
140	115	120	125	
141	Glu Phe Glu Ala Thr Tyr Met Gly Leu Lys Gly Ala Gly Val Gly			
142	130	135	140	
143	Gly Thr Thr Gln Leu Gly Lys Asp Asp Gly Asp Glu Ser Ala Ala Glu			
144	145	150	155	160
145	Val Met Met Asp Val Ser Asp Leu Pro Glu Ser Phe Asp Trp Arg Glu			
146	165	170	175	
147	Lys Gly Ala Val Thr Glu Val Lys Thr Gln Gly Arg Cys Gly Ser Cys			
148	180	185	190	
149	Trp Ala Phe Ser Thr Thr Gly Ala Ile Glu Gly Ala Asn Phe Ile Ala			
150	195	200	205	
151	Thr Gly Lys Leu Leu Ser Leu Ser Glu Gln Gln Leu Val Asp Cys Asp			
152	210	215	220	
153	His Met Cys Asp Leu Lys Glu Lys Asp Asp Cys Asp Asp Gly Cys Ser			
154	225	230	235	240
155	Gly Gly Leu Met Thr Thr Ala Phe Asn Tyr Leu Ile Glu Ala Gly Gly			
156	245	250	255	
157	Ile Glu Glu Glu Val Thr Tyr Pro Tyr Thr Gly Lys Arg Gly Glu Cys			

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/559,986

DATE: 09/20/2006  
TIME: 10:53:24

Input Set : E:\PAT 60100W-1 Sequence Listing.txt  
Output Set: N:\CRF4\09202006\J559986.raw

158	260	265	270
159	Lys Phe Asn Pro Glu Lys Val Ala Val Lys Val Arg Asn Phe Ala Lys		
160	275	280	285
161	Ile Pro Glu Asp Glu Ser Gln Ile Ala Ala Asn Val Val His Asn Gly		
162	290	295	300
163	Pro Leu Ala Ile Gly Leu Asn Ala Val Phe Met Gln Thr Tyr Ile Gly		
164	305	310	315
165	Gly Val Ser Cys Pro Leu Ile Cys Asp Lys Lys Arg Ile Asn His Gly		
166	325	330	335
167	Val Leu Leu Val Gly Tyr Gly Ser Arg Gly Phe Ser Ile Leu Arg Leu		
168	340	345	350
169	Gly Tyr Lys Pro Tyr Trp Ile Ile Lys Asn Ser Trp Gly Lys Arg Trp		
170	355	360	365
171	Gly Glu His Gly Cys Tyr Arg Leu Cys Arg Gly His Asn Met Cys Gly		
172	370	375	380
173	Met Ser Thr Met Val Ser Ala Val Val Thr Gln Thr Ser		
174	385	390	395
177	<210> SEQ ID NO: 3		
178	<211> LENGTH: 726		
179	<212> TYPE: DNA		
180	<213> ORGANISM: Coffea canephora		
182	<220> FEATURE:		
183	<221> NAME/KEY: mRNA		
184	<222> LOCATION: (1)..(726)		
185	<223> OTHER INFORMATION: mRNA		
187	<220> FEATURE:		
188	<221> NAME/KEY: CDS		
189	<222> LOCATION: (79)..(498)		
190	<223> OTHER INFORMATION: CDS		
192	<400> SEQUENCE: 3		
193	ggcgcaacaa acattgaaag aaaatcaaga acccaaaaaa accccacaag aaaaaaagaa	60	
194	aaagaagaag aaaagcca atg gca aaa cca tcg tca tct cta ctc aca ctt	111	
195	Met Ala Lys Pro Ser Ser Ser Leu Leu Thr Leu		
196	1 5 10		
197	cct tcc ttt ctt ctg atc ttt ttc att ctt gca cta ttt tcc acc acc	159	
198	Pro Ser Phe Leu Leu Ile Phe Phe Ile Leu Ala Leu Phe Ser Thr Thr		
199	15 20 25		
200	ctc caa gtt aat gcc ttg gga agg aaa gtg gga gca agg gag aag att	207	
201	Leu Gln Val Asn Ala Leu Gly Arg Lys Val Gly Ala Arg Glu Lys Ile		
202	30 35 40		
203	gag gat gtg aag agc aac aaa gaa gtt caa gaa ctt ggg gaa tat tgt	255	
204	Glu Asp Val Lys Ser Asn Lys Glu Val Gln Glu Leu Gly Glu Tyr Cys		
205	45 50 55		
206	gtt tct gag tac aac aag agt ttg cgg aag aag aac gaa agt ggt	303	
207	Val Ser Glu Tyr Asn Lys Ser Leu Arg Lys Lys Asn Asn Glu Ser Gly		
208	60 65 70 75		
209	gct cct ata atc ttc aca tct gtg gtg gag gct gag aag cag gtg gtt	351	
210	Ala Pro Ile Ile Phe Thr Ser Val Val Glu Ala Glu Lys Gln Val Val		
211	80 85 90		

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/559,986

DATE: 09/20/2006  
TIME: 10:53:24

Input Set : E:\PAT 60100W-1 Sequence Listing.txt  
Output Set: N:\CRF4\09202006\J559986.raw

212	gct	ggg	atc	aaa	tat	tat	ctc	aag	att	aag	gcc	acc	act	tct	tct	ggg	399
213	Ala	Gly	Ile	Lys	Tyr	Tyr	Leu	Lys	Ile	Lys	Ala	Thr	Thr	Ser	Ser	Gly	
214				95					100						105		
215	gtt	ccc	aag	gtt	tac	gat	gcc	att	gtg	gtg	gtt	cgg	cct	tgg	gtt	cat	447
216	Val	Pro	Lys	Val	Tyr	Asp	Ala	Ile	Val	Val	Val	Arg	Pro	Trp	Val	His	
217				110				115				120					
218	act	aag	cca	agg	cag	ttg	ctc	aag	ttc	tcc	cct	tcc	cct	gcc	act	aaa	495
219	Thr	Lys	Pro	Arg	Gln	Leu	Leu	Asn	Phe	Ser	Pro	Ser	Pro	Ala	Thr	Lys	
220						125			130			135					
221	tga	agaagaaaaat	gttgaaaaaag	ttggaaactgt	ttgggagatc	taatctgatg											548
222	attatttagta	cctttcagtg	caaattctct	ttgcgtttaa	gtgttcgtt	tttttttttt											608
223	ccctgtgtct	atttatgacc	gtggcatga	tgatatggtg	tatgatccag	taataattaa											668
224	aatctgttgc	ataaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa											726
227	<210>	SEQ ID NO:	4														
228	<211>	LENGTH:	139														
229	<212>	TYPE:	PRT														
230	<213>	ORGANISM:	Coffea canephora														
232	<400>	SEQUENCE:	4														
233	Met	Ala	Lys	Pro	Ser	Ser	Ser	Ser	Leu	Leu	Thr	Leu	Pro	Ser	Phe	Leu	Leu
234	1					5				10						15	
235	Ile	Phe	Phe	Ile	Leu	Ala	Leu	Phe	Ser	Thr	Thr	Leu	Gln	Val	Asn	Ala	
236						20				25			30				
237	Leu	Gly	Arg	Lys	Val	Gly	Ala	Arg	Glu	Lys	Ile	Glu	Asp	Val	Lys	Ser	
238						35			40			45					
239	Asn	Lys	Glu	Val	Gln	Glu	Leu	Gly	Glu	Tyr	Cys	Val	Ser	Glu	Tyr	Asn	
240						50			55			60					
241	Lys	Ser	Leu	Arg	Lys	Lys	Asn	Asn	Glu	Ser	Gly	Ala	Pro	Ile	Ile	Phe	
242							65		70			75			80		
243	Thr	Ser	Val	Val	Glu	Ala	Glu	Lys	Gln	Val	Val	Ala	Gly	Ile	Lys	Tyr	
244							85			90			95				
245	Tyr	Leu	Lys	Ile	Lys	Ala	Thr	Thr	Ser	Ser	Gly	Val	Pro	Lys	Val	Tyr	
246							100			105			110				
247	Asp	Ala	Ile	Val	Val	Val	Arg	Pro	Trp	Val	His	Thr	Lys	Pro	Arg	Gln	
248							115			120			125				
249	Leu	Leu	Asn	Phe	Ser	Pro	Ser	Pro	Ala	Thr	Lys						
250						130			135								
253	<210>	SEQ ID NO:	5														
254	<211>	LENGTH:	2282														
255	<212>	TYPE:	DNA														
256	<213>	ORGANISM:	Coffea canephora														
258	<220>	FEATURE:															
259	<221>	NAME/KEY:	mRNA														
260	<222>	LOCATION:	(1)..(2282)														
261	<223>	OTHER INFORMATION:	mRNA														
263	<220>	FEATURE:															
264	<221>	NAME/KEY:	CDS														
265	<222>	LOCATION:	(439)..(1731)														
266	<223>	OTHER INFORMATION:	CDS														
268	<400>	SEQUENCE:	5														

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/559,986

DATE: 09/20/2006

TIME: 10:53:25

Input Set : E:\PAT 60100W-1 Sequence Listing.txt

Output Set: N:\CRF4\09202006\J559986.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date